ELSEVIER Journal of Computational and Applied Mathematics 89 (1998) 219–223

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

One step semi-explicit methods based on the Cayley transform for solving isospectral flows [1]

F. Diele [a,*], L. Lopez [a], T. Politi [b]

[a] *Istituto per Ricerche di Matematica Applicata C.N.R., Via Amendola 122/I, I-70126 Bari, Italy*

[b] *Dipartimento Interuniversitario di Matematica, Università degli Studi e Politecnico di Bari, Via E. Orabona 4, I-70125 Bari, Italy*

Received 28 October 1997

Abstract

This note deals with the numerical solution of the matrix differential system

$$Y' = [B(t,Y),Y], \quad Y(0) = Y_0, \ t \geqslant 0, \tag{1}$$

where Y_0 is a real constant symmetric matrix, B maps symmetric into skew-symmetric matrices, and $[B(t,Y),Y]$ is the Lie bracket commutator of $B(t,Y)$ and Y, i.e. $[B(t,Y),Y] = B(t,Y)Y - YB(t,Y)$. The unique solution of (1) is isospectral, that is the matrix $Y(t)$ preserves the eigenvalues of Y_0 and is symmetric for all t (see [1, 5]). Isospectral methods exploit the Flaschka formulation of (1) in which $Y(t)$ is written as $Y(t) = U(t)Y_0U^{\mathrm{T}}(t)$, for $t \geqslant 0$, where $U(t)$ is the orthogonal solution of the differential system

$$U' = B(t, UY_0U^T)U, \quad U(0) = I, \ t \geqslant 0, \tag{2}$$

(see [5]). Here a numerical procedure based on the Cayley transform is proposed and compared with known isospectral methods. © 1998 Elsevier Science B.V. All rights reserved.

Keywords: Cayley transform; Isospectral flows

AMS classification: 65L05; 65L06

1. The method of the Cayley transform

In the following, we shall use these notations: $\mathcal{M}_n$ is the set of $n \times n$ real matrices, $\mathcal{S}_n$ is the subset of the symmetric matrices, $\mathcal{O}_n$ is the subset of the orthogonal matrices, $\mathcal{H}_n$ is the subset of

* Corresponding author. E-mail: pptt@pascal.dm.uniba.it.

[1] This work is partially supported by CNR contract 96.00240.CT01, and Progetto Strategico "Modelli e Metodi per la Matematica e l'Ingegneria" and it has been presented in extended version at NUMDIFF 8, Alexisbad (Germany), 1–5 September 1997.

PII S0377-0427(97)00236-7

the skew-symmetric matrices, $\sigma(U)$ is the spectrum of the matrix U, $C^1[S_1 \to S_2]$ is the set of continuous and differentiable matrix functions mapping S_1 to S_2. We introduce numerical methods based on the Cayley transform (see [6, p. 73]): let us consider $A \in C^1[\mathbb{R} \to \mathcal{H}_n]$, then the matrix function

$$U(t) = [I - A(t)]^{-1}[I + A(t)], \quad t \geqslant 0 \tag{3}$$

belongs to $C^1[\mathbb{R} \to \mathcal{O}_n]$. Conversely, if $U \in C^1[\mathbb{R} \to \mathcal{O}_n]$ and if $-1 \notin \sigma(U(t))$ for each t, then there exists a unique solution $A \in C^1[\mathbb{R} \to \mathcal{H}_n]$ such that (3) is verified with

$$A(t) = [U(t) - I][U(t) + I]^{-1} \quad t \geqslant 0. \tag{4}$$

Since $A(t)$ is skew-symmetric its eigenvalues are pure imaginary and matrix $I - A(t)$ is guaranteed to be nonsingular. It is possible to prove that if $U(t)$ is the orthogonal solution of (2) on $[0,\tau)$, with τ sufficiently small, then matrix function $A(t)$ defined in (4) is well defined for each $t \in [0,\tau)$ and satisfies the following skew-symmetric differential system

$$A' = H(t_k + t, A), \quad A(0) = 0, \ t \in [0,\tau), \tag{5}$$

with $H: \mathbb{R} \times \mathcal{H}_n \to \mathcal{H}_n$, given by $H(t_k + t, A) = \frac{1}{2}(I - A)B(t_k + t, UY_kU^T)(I + A)$.

The advantage of this procedure is that all classical numerical methods are skew-symmetric preserving integrators, thus explicit methods may be employed to reduce the cost of the numerical procedure. We consider the v-stage explicit RK method defined by the matrix $\boldsymbol{A} = (a_{ij})$ and the vectors $\boldsymbol{c} = (c_1, \dots, c_v)^{\mathrm{T}}$ and $\boldsymbol{b} = (b_1, \dots, b_v)^{\mathrm{T}}$. Then the approximation A_1 of the solution of (5) at $t = h$ is given by

$$A_1 = h \sum_{l=1}^{v} b_l H(t_k + c_l h, A_{1l}), \qquad A_{1l} = h \sum_{j=1}^{l-1} a_{lj} H(t_k + c_j h, A_{1j}), \quad l = 2, \dots, v, \tag{6}$$

with $A_{11} = 0$, $H(t_k + c_l h, A_{1l}) = \frac{1}{2}(I - A_{1l})B(t_k + c_l h, U_{1l} Y_k U_{1l}^T)(I + A_{1l})$ and

$$(I - A_{1l})U_{1l} = (I + A_{1l}) \quad l = 1, \dots, v. \tag{7}$$

Once A_1 has been derived, U_1 may be computed by solving the linear matrix system

$$(I - A_1)U_1 = (I + A_1) \tag{8}$$

and the isospectral correction

$$Y_{k+1} = U_1 Y_k U_1^{\mathrm{T}}, \tag{9}$$

may be performed. The procedure stated in (6)–(9) will be denoted by CAYRKv.

Theorem 1.1. *Let us consider the isospectral system* (1) *and let* $B: \mathbb{R} \times \mathcal{S}_n \to \mathcal{H}_n$ *be sufficiently smooth. Then the CAYRKv method* (6)–(9) *has the same order of accuracy of the underlying v-stage explicit RK method.*

Proof. We consider a CAYRKv method based on an explicit RK scheme of order p such that the local truncation error is $A_1 - A(h) = O(h^{p+1})$. By using (7) and the Cayley transform for the exact solution $A(h)$, we have

$$(I - A_1)U_1 - (I - A(h))U(h) = A_1 - A(h),$$

from which it follows that

$$U_1 - U(h) = (I - A(h))^{-1}(A_1 - A(h))(I + U_1).$$

Since $A(h)$ is skew-symmetric, it is also normal together with $(I - A(h))^{-1}$ then it follows that $\|(I-A(h))^{-1}\| \leqslant 1$, where $\|\cdot\|$ is the Euclidean norm on matrices. Furthermore, since U_1 is orthogonal, we have

$$\|U_1 - U(h)\| \leqslant 2\|A_1 - A(h)\|,$$

therefore, $U_1 - U(h)$ behaves as $A_1 - A(h)$. If we assume the local condition $Y(kh) = Y_k$, it follows that

$$Y_{k+1} - Y((k+1)h) = [U_1 - U(h)]Y(kh)U_1^{\mathrm{T}} - Y((k+1)h)[U_1 - U(h)]U_1^{\mathrm{T}}.$$

Thus, from the isospectrality of the theoretical solution, it follows that $\|Y(kh)\| = \|Y_0\|$ for all $k \geqslant 0$, therefore we obtain

$$\|Y_{k+1} - Y((k+1)h)\| \leqslant 2\|Y(kh)\|\|U_1 - U(h)\| \leqslant 4\|Y_0\|\|A_1 - A(h)\|,$$

that is the local truncation error $Y_{k+1} - Y((k+1)h)$ behaves as the local truncation error $A_1 - A(h)$, that is as $O(\|Y_0\|h^{p+1})$. $\square$

2. Computational aspects and numerical tests

Here, given a grid point t_k, the computational costs (in terms of flops) of some semi-explicit isospectral methods are compared. For simplicity, we neglect the computational costs of order lower than n^3, while we consider the cost for multiplying matrices, the cost for solving linear systems, the cost $\alpha(n)$ for computing the Lie-bracket operator, the cost $\delta(n)$ for computing $B(t, Y)$. We denote by ADJRKv the method of the adjoint equation (proposed in [2]) in which the Flaschka and adjoint equation are both solved by a v-stage explicit RK method. We denote by PRKv the semi-explicit method in which U_1 is computed by a projection method based on the modified Gram–Schmidt algorithm (see [4]). We denote by EvGLRKs the semi-explict methods of order $2s$ based on v-stage continuous explicit RK methods. We recall that v depends on s, in particular, we have $v(1) = 1$, $v(2) = 4$, $v(3) = 7$ (see [7]). We note that, the flop count for multiplying the matrices in (9) is almost $3n^3$ being Y_{k+1} symmetric. Observe that if $Y \in \mathcal{S}_n$, then $[B(t,Y), Y] = B(t,Y)Y + (B(t,Y)Y)^{\mathrm{T}}$, thus $\alpha(n) \cong \delta(n) + 2n^3$. For instance, in the case of Toda and double bracket flows (see examples 2.1 and 2.3), we get, respectively, $\alpha(n) = O(2n^3)$ and $\alpha(n) = O(4n^3)$ flops. Furthermore, if $Y \in \mathcal{S}_n$ and $A \in \mathcal{H}_n$, then

$$2H(t,A) = (I - A)B(t,Y)(I + A) = B(t,Y) + [B(t,Y),A] - AB(t,Y)A,$$

Table 1

Method	Flop Count
EvGLRKs	$v(\delta(n)+2n^3)+[s\delta(n)+(\frac{2}{3}s^3+2s^2)n^3+(2s+3)n^3]$
PRKv	$v(\delta(n)+2n^3+6n^3)-n^3$
CAYRKv	$v(\delta(n)+\frac{8}{3}n^3+6n^3)-3n^3$
ADJRKv	$v(\delta(n)+2n^3+6n^3)-2n^3$

therefore, since $AB(t,Y)A$ is skew-symmetric, it follows that the computation of $H(t,A)$ requires $\delta(n)+3n^3$ flops. We suppose to use algorithms based on Gaussian elimination to solve linear systems and to invert matrices. In particular, we suppose that the cost for solving n linear systems with the same coefficient matrix of dimension sn is given by $(\frac{2}{3}s^3+2s^2)n^3$ flops. Table 1 may thus be derived.

From this Table 1 it follows that, among second-order methods, E1GLRK1 seems to be the less expensive procedure while the other methods show similar flop counts. Notice that this kind of cost analysis is only a rough guide because to evaluate the performance of a numerical method the flop count (or the CPU time) should be related to the global errors at the grid point. In the sequel, comparisons among the performances of the studied methods will be reported. All the numerical results have been obtained by Matlab codes implemented on a scalar computer Alpha 200 4/233 with 128 Mb RAM. The following examples, we have considered have already been described in [7].

Example 2.1. (*Toda flow*). Let us consider the isospectral flow (1) in which $B(t,Y)=Y_+-Y_-$, where Y_+ (Y_-) denotes the upper (lower) triangular part of Y.

Example 2.2. (*Continuous power method*). In this case $B(t,Y)=T\circ Y$, where $\circ$ denotes the Hadamard product and $T=[t_{ij}]$ is the skew-symmetric matrix such that $t_{1,i}=-t_{i,1}=1$, for $i>1$ and with all other elements equal to 0.

Example 2.3. (*Double-bracket flow*). Let us consider the isospectral flow (1) in which $B(t,Y)=[D,Y]$, with $D=\mathrm{diag}(1,0.8,0.6,0.4,0.2)$.

In Figs. 1.1, 1.2 and 1.3 we have plotted the performances at $t=15$ of the second-order methods considered applied, respectively, to examples 2.1, 2.2 and 2.3, with steps $h=1/2^i$, for $i=2,3,4,5,6,7$, solid line represents CAYRK2, dash-dotted line ADJRK2 and dashed line E1GLRK1. The E1GLRK1 require always less flops but give a higher error for Example 2.1, while in the other two examples the best performance is given by CAYRK2. In Fig. 1.4 we have plotted the symmetric error of the numerical solution Y_k, computed using $\|Y_k-Y_k^{\mathrm{T}}\|_F$, for Example 2.2. In this case the limit matrix computed by ADJRK2 is not symmetric and probably this is the reason of higher errors. The symmetric error given by E1GLRK1 and CAYRK2 is negligible. This behaviour has not been noted in Example 2.1 since, in this case, the limit matrix is a diagonal matrix.

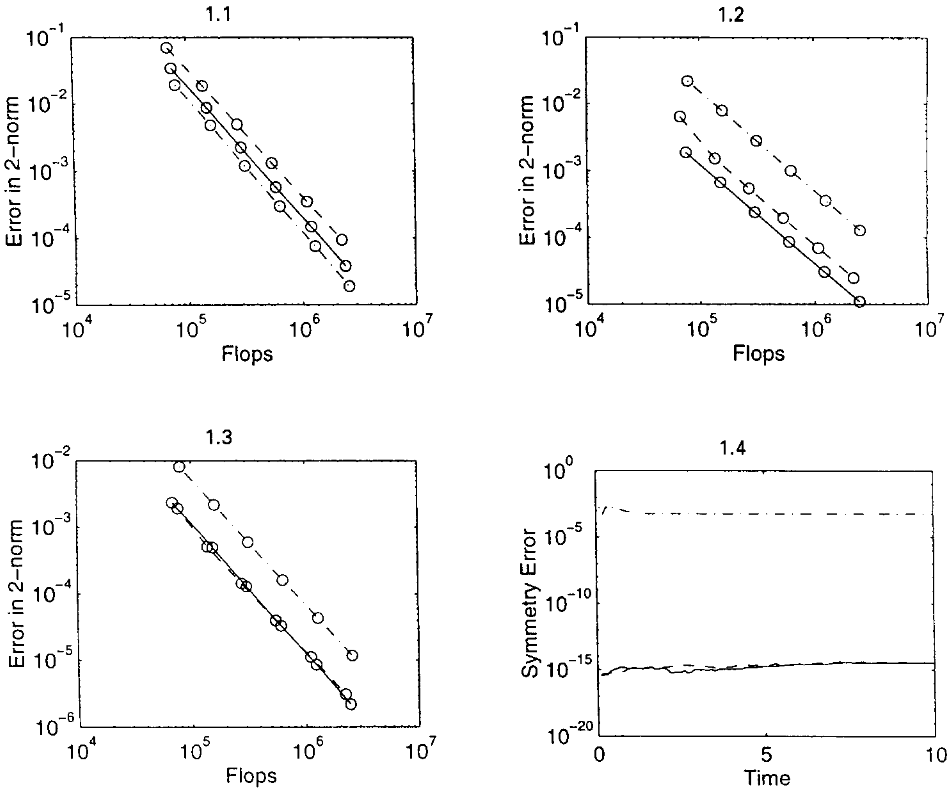

Fig. 1.

References

[1] M. Calvo, A. Iserles, A. Zanna, Numerical solution of isospectral flows, Math. Comp. to appear.

[2] M. Calvo, A. Iserles, A. Zanna, Runge–Kutta methods for orthogonal and isospectral flows, Appl. Numer. Math. 22 (1996) 153–164.

[3] F. Diele, L. Lopez, R. Peluso, The Cayley transform in the numerical solution of unitary differential systems, IRMA-CNR Report 5/96, Adv. Comput. Math. submitted.

[4] L. Dieci, D. Russell, F. Van Vleck, Unitary integration and applications to continuous orthonormalization techniques, SIAM J. Numer. Anal. 31 (1994) 261–281.

[5] H. Flaschka, The Toda lattice, I. Phys. Rev. B 9 (1974) 1924–1925.

[6] G. Golub, C. Van Loan, Matrix Computations, Johns Hopkins University Press, Baltimore, Maryland, 1996.

[7] L. Lopez, T. Politi, Numerical procedures based on Runge–Kutta methods for solving isospectral flows, App. Numer. Math. 25 (1997) 1–17.